ELSEVIER
Topology and its Applications 99 (1999) 237–261
TOPOLOGY AND ITS APPLICATIONS
www.elsevier.com/locate/topol

The classification of metrics and multivariate statistical analysis ☆

Stephen Watson [1]

Department of Mathematics, York University, 4700 Keele St., North York, Ontario, Canada M3J 1P3

Received 16 October 1996; received in revised form 4 July 1997

Abstract

We prove some new results on the classification of metrics. These types of metrics include the ultrametrics, the additive metrics, the L_p-embeddable metrics, the hypermetrics, the metrics of negative type and the metrics with one positive eigenvalue. We pose many open questions.

Keywords: Ultrametric; Isometric; Hypermetric

AMS classification: Primary 54E35; 51K05, Secondary 62H30; 62H11; 05C12; 68Q15

1. Introduction

Everyone is familiar with the *triangle inequality*. This inequality played a major role in the definition of a topological space.

$$\rho(a, b) \leqslant \rho(a, c) + \rho(b, c).$$

Still familiar to topologists is the *ultrametric inequality*.

$$\rho(a, b) \leqslant \max\{\rho(a, c), \rho(b, c)\}.$$

But there are more inequalities of importance to mathematics which topologists are not familiar with. For example, there is the *four-point inequality*,

$$\rho(a, b) + \rho(c, d) \leqslant \max\{\rho(b, c) + \rho(a, d), \rho(a, c) + \rho(b, d)\}$$

and there is the *pentagon inequality*

☆ This paper was intended as part of the special issue devoted to the 8th Prague Topological Symposium (Volume 85, 22 May 1998). We apologize for the delay in publication. This work has been supported by the Natural Sciences and Engineering Research Council of Canada.
[1] E-mail: watson@msfacb.math.york.ca.

PII: S0166-8641(99)00157-1

$$\rho(a,b)+\rho(c,d)+\rho(c,e)+\rho(d,e)$$
$$\leqslant \rho(a,c)+\rho(a,d)+\rho(a,e)+\rho(b,c)+\rho(b,d)+\rho(b,e),$$

and there is the *negative-type inequality*

$$\rho(a,b)+\rho(b,c)+\rho(a,c)+\rho(d,e)+\rho(d,f)+\rho(e,f)$$
$$\leqslant \rho(a,d)+\rho(a,e)+\rho(a,f)+\rho(b,d)$$
$$+\rho(b,e)+\rho(b,f)+\rho(c,d)+\rho(c,e)+\rho(c,f).$$

All of these inequalities turn out to be important in various parts of mathematics and especially in the applications of mathematics to the sciences.

2. Statistics

The standard definition states that multivariate statistical analysis and, especially, that more applied part of multivariate statistical analysis which is called multivariate data analysis, is concerned with data collected on several dimensions of the same individual. A cursory examination of the literature of that subject reveals that a major concern, worthy of a few chapters in a typical textbook, is the following situation and resulting problem: For each of n objects, each of k tests is performed with a result which might be a real number. This gives us an $n \times k$ matrix. We wish to combine these test data and produce an $n \times n$ matrix of non-negative reals which measures the "similarity" or "dissimilarity" of the objects so far as their test results indicate. If the tests have been designed to give a reasonable notion of similarity, then this *similarity matrix* usually satisfies the axioms of a metric space. We wish to determine what kind of distance concept has been isolated, that is, what kind of metric space has been constructed. [2] Of course, with real data, things are not as simple as I have described. In most cases, the data has to be *transformed*, some data is missing and has to be *reconstructed* and the data has error or even spurious entries and has to be *approximated*. Only then can the data be *represented* in some fashion which makes it possible to use our human facilities to understand this data.

So before analyzing distance data, we need some means of *classifying* metric spaces and some compendium of reasonable representations or *embeddings*.

3. Kinds of metrics

Here is a list of the basic kinds of metrics:

(1) ultrametric,
(2a) L_2-embeddable,
(2b) four-point property,
(3) L_1-embeddable,
(4) hypermetric,
(5) spherical,

[2] This topic is a huge one. There are many textbooks devoted to the various aspects of this problem. A bibliography listing only articles which appeared up to 1975 has 7530 entries.

(6) negative-type,
(7) one positive eigenvalue,
(8) L_∞-embeddable.

Each property implies those properties listed below it, except that (2a) does not imply (2b).

The purpose of this article is to introduce these classes of metrics, to describe their importance to mathematics and the sciences, to state the basic theorems concerning these classes, to state some new theorems which I have obtained using topological methods, and even provide a proof here and there. But the main purpose is to state many of the open problems around these concepts and to show you how much of this subject might be understood by topological means.

4. Ultrametrics

Ultrametric spaces are well known to topologists and perhaps even better known to number theorists and analysts. Hensel invented the p-adic numbers in 1897. These numbers carry a natural ultrametric structure and there are now textbooks on “Ultrametric Calculus” and “Non-Archimedean Functional Analysis”. A closely related topic which has attracted attention of many topologists is spherical completeness. The ultrametric inequality was formulated at least as early as 1934 by Hausdorff but the term *ultrametric* was coined only in 1944 by Krasner. In 1956 deGroot characterized the ultrametric spaces, up to homeomorphism, as the strongly zero-dimensional metric spaces. [3]

But questions at a topological level of generality remain open. It does not seem to be known which non-metric spaces are “essentially” ultrametric:

Problem 1. Characterize those topological spaces X such that, for every continuous pseudometric ρ on X, there is a continuous ultra-pseudometric σ on X which generates a larger topology.

Ultrametric spaces have emerged in the last fifteen years as a major concern in statistical mechanics, in neural networks and in optimization theory. The history of this emergence is quite interesting.

In 1984, Mezard, Parisi, Sourlas, Toulouse and Virasoro published an article on the mean-field theory of spin glasses in which they established that the distribution of “pure states” in the configuration space is an ultrametric subspace. Within a few years, it was shown that the “graph partitioning problem” in finite combinatorics could be “mapped onto” the spin glass problem and thus that the solution space for this problem also has an ultrametric structure. Kirkpatrick then found numerically that the solutions for certain *traveling salesman problems* [4] seem to scatter in an ultrametric fashion. Bouchaud and

---

[3] Nyikos and Purisch have extensively investigated the relationship between ultrametrics and generalized metrics and orderability.

[4] Is there an infinitary version of the traveling salesman problem? Examples might be “When do metric spaces admit space-filling curves of finite length?” or “When do they admit ε-dense curves of finite length for each $\varepsilon > 0$?”

Le Doussal [14] have conjectured that, in optimization problems in which "the imposed constraints cannot all be satisfied simultaneously, the optimal configurations (i.e., those which minimize the number of unsatisfied constraints) spread in an ultrametric way in the configuration space". These kinds of problems are known as *frustrated optimization problems*.

No results of this kind have actually been proven except in special classes of spin glasses. All other indications are numerical or by reduction. It would be of major significance to many fields to show that this phenomenon occurs under some general circumstances.

Problem 2. Give some reasonable conditions on non-negative continuous real-valued functions $\{f_i\colon\ i < n\}$ on a metric space X so that, if K is minimal for $Y = \{x \in X\colon \sum\{f_i(x)\colon\ i < n\} = K\}$ non-empty, then Y is ultrametric. Formulate this question more accurately.

A recent and effective strategy in handling optimization problems is to use simulated annealing and *random walks* to find global solutions. In problems where the local solutions have an ultrametric structure, it is therefore essential to understand random walks on ultrametric spaces. There has been much work already on different ways in which to define such random walks.

There are, undoubtedly, quite general theorems which show that the natural metric on sufficiently few independent stochastic processes which are nontrivial on sufficiently few of sufficiently many coordinates is arbitrarily close to being ultrametric. It seems likely that, to obtain a statement and proof of such a theorem, we should state and prove an infinite version first.

Problem 3. Let $\{R^x\colon\ x \in X\}$ be a finite set of independent stochastic processes acting on $\mathbb{R}^\omega$, independently of the coordinates, so that

$$(\forall x \in X)(\forall t \in \mathbb{R})\ \mathrm{Prob}_t\big(|\{n \in \omega\colon\ R^x(n) \neq 0\}| < \omega\big) = 1.$$

Let d be the metric defined on X by

$$d(x, x') = E\big(L_1(|R^x - R^{x'}|)\big)$$

for a suitable measure on ω. Prove that d is an ultrametric.

Problem 4. Can an asymptotic finitary version of Problem 3 be stated and proved? Can the assumptions be made sufficiently reasonable so as to show that the numerical evidence for ultrametricity of phylogenetic trees in evolution is inevitable?

In examining the numerical evidence for ultrametricity, and in proving theoretical results about the tendency of finite data to approach ultrametricity, there is a need for answering a fundamental question: How can we measure how far a given metric is from being an ultrametric?

The main method used in spin glasses for answering this question is based on the following:

Proposition 1 (Jardine, 1967). *If ρ is a metric on a finite set, then there is an ultrametric τ which minimizes* $\sup\{|\rho(x,y)-\tau(x,y)|\colon x,y\in X\}$ *among those τ such that* $(\forall x,y\in X)$ $\tau(x,y)\leqslant\rho(x,y)$.

This analog of the subharmonic in potential theory which is called the subdominant ultrametric can be quite pathological. Rammal, Toulouse and Virasoro in their article "Ultrametricity for Physicists" [55] ask whether there are optimal l_p ultrametric approximations for a given metric where $1\leqslant p\leqslant\infty$ (and specifically ask it for 1 and ∞). Noting that the proposition can be viewed as an optimal l_∞ ultrametric approximation among those ultrametrics below a given metric, we have obtained the next result:

Theorem 1. *If ρ is a metric on a finite space, then there is an ultrametric τ which minimizes* $\sup\{|\rho(x,y)-\tau(x,y)|\colon x,y\in X\}$ *among all ultrametrics τ.*

There may be several choices for the ultrametric in Theorem 1 but perhaps this duplication only occurs in a trivial way.

Problem 5. Is there, up to some kind of manipulation, always a unique ultrametric τ which minimizes $\sup\{|\rho(x,y)-\tau(x,y)|\colon x,y\in X\}$ among all ultrametrics τ?

But our construction in Theorem 1 seems to take exponential time while Jardine's only takes polynomial time.

Problem 6. Is there a polynomial algorithm for computing an ultrametric τ which minimizes $\sup\{|\rho(x,y)-\tau(x,y)|\colon x,y\in X\}$ among all ultrametrics τ?

Krivanek showed that computing the closest ultrametric above a given metric is NP-complete.

Problem 7. Show that the subdominant ultrametric can be quite pathological. That is, show that the subdominant ultrametric of a given metric ρ can be arbitrarily close to zero, even when there is an ultrametric quite close to ρ in the supremum norm.

Jardine's theorem was extended by Bayod and Martinez-Maurica in 1990 [12] to totally disconnected locally compact spaces. But they failed to obtain a characterization.

Problem 8. Characterize those metric spaces which have a subdominant ultrametric.

Problem 9. Can Theorem 1 be extended to a reasonable class of infinite metric spaces?

Returning to the problem of Rammal, Toulouse and Virasoro [55]:

Problem 10. If ρ is a metric on a finite (or arbitrary) set, then is there an ultrametric τ which minimizes $\sum\{|\rho(x,y)-\tau(x,y)|\colon x,y\in X\}$ among all ultrametrics τ? How does one construct τ?

Problem 11. Which metric spaces have a (uniformly) equivalent metric ρ for which there is an ultrametric τ such that $\sum\{|\rho(x,y)-\tau(x,y)|\colon x,y\in X\}$ is finite?

It would be quite useful to associate, to each metric, an ultrametric which is somehow derived from it in a natural way. But this seems unlikely.

Problem 12. Let the family of all metrics on a (finite, countable or arbitrary) set X be equipped with an l_p metric. Is there a continuous retraction of metrics onto ultrametrics?

Note that when $p=\infty$, this problem is entirely topological.

Ultrametric spaces can be embedded in linearly ordered spaces but this is not an isometric embedding. To provide an isometric representation we must use another device, well known to natural scientists as a *dendrogram* (see [55, p. 769]). This method is equally valid for infinite spaces.

5. Additive trees

The representation of ultrametrics by dendrograms leads one to consider a more general kind of diagram called an additive tree in the social sciences literature or a phylogenetic tree (this term has many inexact definitions) in the biological literature. Suppose (V,E) is a tree (a graph without cycles or loops) in which each edge has a "weight" which is a non-negative real number. The distance between any two vertices $x,y\in V$ is defined to be the sum of the weights of the edges which make up the unique minimal path from x to y. It is an exercise in graph theory to show that this distance is a metric which satisfies the four-point property.

Theorem 2. *Any ultrametric space satisfies the four-point property.*

In 1971, Bunemann showed that, in fact, any metric on a finite set satisfying the four-point property could be represented as the vertices of a graph equipped with this "path distance".

Definition 1. An *R-tree* is a (uniquely) arcwise connected metric space in which each arc is isometric to a subarc of the reals.

In 1985, Mayer and Oversteegen constructed a universal R-tree of a given weight. This construction allows us to prove that the path metric or *intrinsic* metric on an R-tree satisfies the four-point condition and that, conversely, any metric space satisfying the four-point condition can be represented as a set of points in a R-tree. [5]

Indeed the concept of an additive tree may be valuable for arbitrary completely regular spaces:

Problem 13. Characterize those topological spaces X such that, for every continuous pseudometric ρ on X, there is a continuous pseudometric σ on X with the four-point property which generates a larger topology.

Any linearly ordered connected compactum satisfies Problem 13.

This representation by additive trees is not, by any means, only a theoretical concern. It is a useful way of representing data which satisfies the four-point condition (see Shepard [59, p. 395]). Note that this is the right diagram for representing evolution in which rates of evolution may be different for different species. Dendrograms assume that the rates are uniform for all species.

Additive trees are obviously easy to interpret. A topologist might ask whether one can use the intrinsic metric of more general spaces to represent metric spaces of a broader kind. The answer is yes.

Proposition 2. *Any separable metric space can be represented as a subset of a subspace of* $\mathbb{R}^3$ *equipped with the intrinsic metric.*

But this proposition shows by its strength, its uselessness. We must keep in mind that, to be useful, a representation must take advantage of human facilities. [6]

Problem 14. Characterize those metric spaces which can be represented as a subset of a (simply connected) continuum in $\mathbb{R}^2$ with the intrinsic metric.

For example, any ultrametric space, such as K_5 with the graph metric, [7] can be so represented but $K(3,3)$ cannot be so represented.

Problem 15. Is there a version of Kuratowski's test for planarity of graphs which answers Problem 14 for graph metrics? That is, is there a finite list of "forbidden" graphs?

While testing a metric for ultrametricity requires testing each set of three points (and thus can be done in $O(n^3)$ computing time), testing a metric for the four-point condition

---

[5] Rudnik and Borsuk have asked whether there is a one-dimensional subset X of $\mathbb{R}^2$ in which every two points is joined by an arc of finite length and in which every intrinsic isometry in $\mathbb{R}^2$ is an isometry.

[6] But, despite this, many articles in the optimization literature ask for minimizing the total length of a graph which represents a given finite metric space. This should also be explored for infinite metric spaces.

[7] Any connected graph has a *graph metric* which is the largest metric in which the distance between any two vertices which are joined by an edge is 1.

seems to require testing each set of four points and that would require $O(n^4)$ time. But there is a beautiful way of converting additive trees into ultrametrics.

Definition 2. If ρ is a metric on a set X and $v \in X$ and c is an appropriate constant, then, for each $x, y \in X$, define $\delta(x, y) = c + \rho(x, y) - \rho(x, v) - \rho(y, v)$. δ is the *Farris transform* of ρ.

Proposition 3 (Farris, 1970). *δ is an ultrametric if and only if ρ satisfies the four-point condition.*

This theorem is not hard to prove: it just requires some manipulation. Of course δ and ρ do not generate the same topology even if we choose c carefully.

But Farris' lemma is quite useful. We see immediately that we can test the four-point condition in just $O(n^3)$ time. Actually testing ultrametricity and thus the four-point condition can even be done in $O(n^2 \log n)$ time.

Problem 16. Which metric spaces can be represented up to uniform equivalence by a subset of a space (or an R-tree) with the intrinsic metric and finite total length?

6. L_1-embeddable metrics and their decompositions

A metric space (X, ρ) where X is finite is said to be l_1-*embeddable* if we can embed X isometrically into l_1.

Do such metric spaces occur in nature? Is this class useful for statistical analysis? It is often true that real-life estimates of similarity are obtained by forming a linear combination of various criteria. Such estimates, such metrics are precisely the L_1-embeddable metrics! Let us make this exact.

Definition 3. Let $(X, \mathcal{M}, \sigma)$ be a measure space. For $A, B \in \mathcal{M}$, define $\rho(A, B) = \int_{A \triangle B} \mathrm{d}\sigma$. We call ρ L_1-*embeddable*.

Since we use integration, we are restricted to estimating similarity by linear combinations of various criteria. But this still allows us to represent a broad range of metrics.

Proposition 4. *Let ρ be a metric on a finite set. Then ρ is l_1-embeddable if and only if ρ is L_1-embeddable.*

Theorem 3. *If a metric ρ on X satisfies the four-point-condition, then ρ is L_1-embeddable.*

Proof. Represent (X, ρ) by a subset of a R-tree Y with the intrinsic metric. Choose $v \in X$. For each $x \in X$, let A_x be the unique shortest path in Y from x to v. Let $\mathcal{M}$ be the set of all Borel sets of Y. Let μ be the measure which assigns to each Borel set B the sum of the lengths of all disjoint families of paths in B. Let f be the constant one function. Now the intrinsic metric between x and y coincides with the L_1 metric on $(Y, \mathcal{M}, \sigma)$. $\square$

In the analysis of statistical data, it is not only important to recognize L_1-embeddable distances but also to be able to decompose distance data into an L_1-combination of more primitive distances. That is, we want to be able to carry out "linear decompositions" whenever this is possible and to identify when this is not possible.

Definition 4. Suppose (X, ρ) is a metric space. If there are metric spaces $\{(X_i, \rho_i)\colon i \in I\}$ and a one-to-one map $\pi : X \to \prod\{X_i\colon i \in I\}$ such that

$$(\forall x, y \in X)\rho(x, y) = \sum_{i \in I} \bigl|\pi(x)(i) - \pi(y)(i)\bigr|$$

and if $\{\pi(x)(i)\colon x \in X\} = X_i$, then we say π is a decomposition (X, ρ) as a subdirect L_1-product of metric spaces.

This is motivated by the important existence of subdirect representations in algebra.

Theorem 4. *Every metric space can be decomposed in a "maximal" manner as a subdirect L_1-product of subsets of the reals and one more irreducible metric space. Every L_1-embeddable metric space is decomposed completely into a subdirect L_1-product of subsets of the reals.*

Proof. Construct π inductively on a well-ordered set I.[8] If this has been done on an initial segment $J \subset I$ and i is the least element of $I - J$, then define

$$\rho^*(x, y) = \rho(x, y) - \sum\bigl\{|\pi(x)(i) - \pi(y)(i)|\colon i \in J\bigr\}$$

and let $\Sigma = \{\sigma \in \mathbb{R}^X\colon\ \rho^* - \sigma$ satisfies the triangle inequality$\}$ be partially ordered by defining $\sigma \leqslant \sigma'$ if, for all $x, x' \in X$, $\sigma(x, x') \leqslant \sigma'(x, x')$. Choose a maximal $\sigma \in \Sigma$ and define, for each $x \in X$, $\pi(x)(i) = \sigma(i)$. $\square$

Problem 17. Is there a "maximal" decomposition of metric spaces as a subdirect L_1-product of additive trees (or Hilbert spaces) and one more irreducible metric space so that every additive tree (or Hilbert space) remains its unique factor?

The notion of L_1-decomposition is well-motivated by the central importance of "dimension reduction" in multivariate data analysis. In his influential textbook, Kshirsagar said "The aim of the statistician undertaking multivariate analysis is to reduce the number of variables by employing suitable linear transformations ... thus reduces the dimensionality of the problem". Reasonable decompositions accomplish this by removing the interaction between coordinates.

Problem 18. Are there reasonable L_p decompositions for $1 < p \leqslant \infty$?

[8] The reals themselves can be decomposed into two copies of the reals, namely as the line $y = x$ and this is why we require an well-ordering of the factors. With a restriction to integer-valued metrics, this is no longer an issue.

A more useful L_1-decomposition would do more and break down the remaining irreducible factor in Theorem 4 into an L_1-product of other irreducible factors whenever possible. I am able neither to prove such a theorem or even to formulate this accurately. The criterion by which such a decomposition should be judged is that it should have as a corollary the following result of Graham and Winkler and reported in [31, p. 7259].

Theorem 5 (Graham and Winkler). *Any finite graph can be canonically embedded isometrically into a maximum Cartesian product of irreducible factors.*

The existence of the decomposition by subdirect products for varieties is a true theorem of universal algebra but this is not a variety and so this seems to be of no help.

The general problem of identifying L_1-embeddability turns out to be significant in operations research. The problem of *multicommodity flows* is set in a graph in which each edge has a capacity and a demand. We seek a flow on the edges of the graph so that flow on each edge meets demand and does not exceed capacity. The so-called Japanese theorem of 1971 states that a capacity and demand are *feasible*, i.e., can be met if there is a metric ρ on the vertices of the graph so that $(c - r)\rho \geqslant 0$. The celebrated Ford–Fulkerson theorem in operations research is just this theorem in the special and tractable case of single commodity flows in which the demand occurs on a single edge. Usually the Ford–Fulkerson condition is not sufficient when the demand is more complicated. However Lomonosov showed in 1978 that this condition is still sufficient when the demand lies on an L_1-embeddable subgraph.

7. Graphs and Hamming distance

Indeed Theorem 5 illustrates the intimate connection between L_1-embeddability and Hamming distance. If we use factors in which all non-zero distances are 1 and a counting measure, then the L_1-distance is precisely the Hamming distance. This Hamming distance is useful in estimating distances between binary strings since error-correcting codes can be designed which do nothing more than replace a string with the "closest" string of a certain kind. Although Avis showed that any finite L_1-embeddable metric space embeds in a "weighted" hypercube, it is not true that an integer-valued L_1-embeddable metric can be embedded in the *hypercube* 2^κ with the Hamming distance.

Problem 19. Give necessary and sufficient conditions for a integer-valued (L_1-embeddable) metric to be embeddable in 2^κ with the Hamming distance.

For example, a necessary condition is that triangles must have even perimeter.

There is a huge literature on graphs which can be embedded in hypercubes and metrics which can be embedded in graphs [9] but this beautiful theory carries us too far away from our topic.

8. Compactness and L_∞-embeddable metrics

A classical result of Banach and Mazur, published in 1932, states that any separable metric space can be isometrically embedded in $L_\infty(\kappa)$ when κ is the continuum. But more is true. Suppose (X, d) is a metric space. Fix $a \in X$ and define an isometric embedding π of X into $C^*(X) \subset L_\infty(|X|)$ by defining $\pi(x)$ by setting $\pi(x)(x') = d(x, x') - d(a, x')$.

Theorem 6 (Banach and Mazur, 1932). *Any metric space can be isometrically embedded in $L_\infty(\kappa)$ for sufficiently large κ.*

This theorem, surprisingly, is essentially finitary.

Theorem 7. *If every finite subset of a metric space X is L_∞-embeddable, then X is L_∞-embeddable.*

Proof. Define, for each finite $F \subset X$, $E(F)$ to be the set of all mappings ϕ from X into $\mathbb{R}^\kappa$ which are isometric when restricted to F and achieve the supremum, for any pair, on a coordinate specifically assigned to that pair. These form a centered family of closed sets. If we restrict ourselves to maps which, for some $x \in X$, satisfy $\phi(x) \equiv 0$, then each $E(F)$ is a subset of a fixed compact set and so we have a nonempty intersection. [10] $\square$

This leads us to the three basic compactness problems for L_∞-embeddable (or L_1-embeddable, or L_p-embeddable) metrics.

- If every finite subset of a metric space X can be embedded in l_∞ (or l_1, or l_p), then must X be embeddable in some L_∞ (or L_1, or L_p)?
- If $n \in \omega$, then what is the minimal $k_n \in \omega$ (if it exists) such that any (l_1-embeddable, l_p-embeddable) finite metric space of size n can be embedded in $l_\infty^{k_n}$ ($l_1^{k_n}$, $l_p^{k_n}$)?
- If $n \in \omega$, then what is the minimal $k_n \leqslant \omega$ (if it exists) such that any metric space which cannot be embedded in l_∞^n (l_1^n, l_p^n) has a subspace of size k_n which also cannot be embedded in l_∞^n (l_1^n, l_p^n)?

For the first of these problems, Witsenhausen showed that, if every finite subset of a metric space X is embeddable in l_1, then X is embeddable in some L_1. Results of Yang and Zhang show that, if every finite subset of a metric space X is embeddable in l_2, then X is embeddable in some L_2. The situation for L_p seems to be unclear:

Problem 20. If every finite subset of a metric space X can be embedded in l_p, then must X be embeddable in some L_p?

[9] Djoković characterized graphs that can be embedded into hypercubes in 1973.

[10] Of course, L_p might not be locally compact but this is irrelevant. We work in the Tychonoff product topology.

Problem 21. Find a general compactness theorem which implies that the solution to the first compactness problem is positive for all p.

For the second problem, Schoenberg noted in 1938 that, although the construction in the proof of Theorem 6 above seems to require n coordinates, we can omit one coordinate without difficulty. This shows that $k_n \leqslant n-1$ for l_∞. Wolfe showed that, in fact, $k_n \leqslant n-2$ for l_∞. Witsenhausen has obtained the lower and upper bounds $n-2 \leqslant k_n \leqslant n(n-1)/2$ for l_1 and, later, Ball showed that $k_n \leqslant n(n-1)/2$ for any l_p. But none of these results solve the problem completely:

Problem 22. If $n \in \omega$, then what is the minimal $k_n \in \omega$ (if it exists) such that any l_1-embeddable finite metric space of size n can be embedded in $l_1^{k_n}$? What about for l_p when $1 < p \leqslant \omega$?

This second problem has an interesting variation. Suppose $D = \{1, 2, 3\}$ has the "distance" in which 1 and 3 are distance one apart and all other pairs are at distance zero. What is the least k_n such that any connected graph on n vertices can be embedded in a product of k_n many copies of D with the L_1 distance? It is not obvious that k_n exists and is finite.

This may seem a strange problem but this is exactly the "addressing problem for loop switching" posed by Graham and Pollak in 1971 in [29] and solved by Winkler in 1983 [66]. The answer is $k_n = n-1$. [11]

The third problem is quite interesting. It may involve finite approximations to topological orientability.

Proposition 5 (Malitz and Malitz, 1992). *If a metric space X cannot be embedded in $\mathbb{R}^2$ with the l_∞-norm* (*or, equivalently, the l_1-norm*)*, then X has a subspace of size* 11 *which cannot be embedded in $\mathbb{R}^2$ with the l_∞-norm* (*or, equivalently, the l_1-norm*)*. Thus determining whether a finite metric space can be embedded in $\mathbb{R}^2$ with the l_∞-norm can be done in polynomial time.*

They state the existence of such a number (like 11), for $\mathbb{R}^n$ when $n \geqslant 3$ is an open question and that their methods get "wildly complicated".

But we have obtained the following results.

Theorem 8. *There is no N such that a finite metric space X cannot be embedded in $\mathbb{R}^3$ with the l_∞-norm if and only if X has a subspace of size N which cannot be embedded in $\mathbb{R}^3$ with the l_∞-norm.*

Proof. Use a Möbius strip in which the width of the strip is much smaller than N times the radius of the circle. Apply compactness to get a finite subset which is still sufficiently "Möbius". □

[11] These are "squashed cubes" but the problem for graphs in ordinary cubes remains open.

Problem 23. Is it true that there is no N such that a finite metric space X cannot be embedded in $\mathbb{R}^3$ with the l_1-norm if and only if X has a subspace of size N which cannot be embedded in $\mathbb{R}^3$ with the l_1-norm? Is this true for some $\mathbb{R}^n$? Can the construction in Theorem 8 be carried out in some $\mathbb{R}^n$ with the l_1-norm?

Theorem 9. *Determining whether a finite metric space X can be embedded in $\mathbb{R}^6$ with the l_∞-norm is* NP-*complete.*

Proof. The axes of a cube can be each be assigned one of three dimensions in exactly six ways. This assignment must be constant on the product of a cube and a line. If we join together two such products in such a way that all coordinates change, then knowing the assignment on one side of the join gives us exactly two possibilities on the other side of the join. Thus using three more dimensions we can code the 3-colorability of graphs which is NP-complete. □

Problem 24. Let $3 \leqslant n \leqslant 5$. Is determining whether a finite metric space X can be embedded in $\mathbb{R}^n$ with the l_∞-norm NP-complete?

Problem 25. Is determining whether a finite metric space X can be embedded in $\mathbb{R}^n$ with the l_1-norm NP-complete?

9. L_2-embeddable metrics

The problem of characterizing metric spaces which embed in Euclidean space of some dimension is a classical one and was solved by Menger in the 1930s. There is a book by Blumenthal entitled *Distance Geometry* and even a *Mathematical Reviews* section 51K devoted to this topic. But, in fact, this is an easy problem in $\mathbb{R}^2$ with the Euclidean (l_2) metric. For if a space embeds in $\mathbb{R}^2$ and a, b, c are points in that space which do not satisfy the equality $\rho(a,b) + \rho(b,c) = \rho(a,c)$ under any permutation, then a is, without loss of generality, embedded arbitrarily. Now b is embedded on some circle centered at a but otherwise its position is arbitrary. We deduce that c must be placed in one of two positions but this choice is again arbitrary. But now any further point must occupy a uniquely determined position. Thus the position of any point is determined uniquely once we have three points "in general position". In the general setting of the Euclidean metric on $\mathbb{R}^n$, the situation is analogous.

Much of the work in distance geometry is devoted to characterizing Euclidean spaces, Banach spaces, hyperbolic spaces, inner product spaces and so forth entirely from the combinatorial properties of their metrics. But we will not discuss here this fascinating topic and its intense activity since 1932 nor will we discuss the interesting work on the "distance-one-preserving" maps of Aleksandrov.

What is surprising and important to us is that ultrametrics are L_2-embeddable.

Theorem 10 (Lemin, 1985; Vestfrid and Timan, 1979, for l_∞). *Any ultrametric space of cardinality κ can be embedded isometrically in generalized Hilbert space*

$$\Big\{f \in \mathbb{R}^\kappa\colon \sum\{f(\alpha)\colon \alpha \in \kappa\} < \infty\Big\}.$$

This requires some work.

Another surprising fact is that L_2-embeddable metrics are L_1-embeddable.

Theorem 11. *Any L_2-embeddable space is L_1-embeddable.*

Problem 26. Give a direct proof that any L_2 embeds isometrically into some $L_1(\mu)$. Can this be done by integration over projections onto hyperplanes of codimension 1? What happens for $p \neq \infty$?

But the most important fact about L_2-embeddable metrics is that they are the basic notion of MDS: *multi-dimensional scaling*. This is a huge topic about which entire books have been written and for which there are many software packages being sold.

The basic purpose of MDS, the thing that these packages accomplish, is to take a set of data, either an $n \times k$ matrix showing the results of tests or an $n \times n$ matrix which already exhibits similarity data, and to do the best job possible in representing this data as points in the plane or in a higher-dimension Euclidean space.

There is a lot involved here. Scaling the similarity data with real numbers, reconstruction of missing and spurious data, approximation to a metric which is embeddable in some Euclidean space. The problem of reconstructing missing data is an important one. Sippl and Scheraga [60] and Schlitter [58] in pursuit of reconstructing distance data in problems on nuclear magnetic resonance showed that we need only a $4 \times n$ submatrix of the distance matrix to reconstruct effectively in $\mathbb{R}^3$ so long as the 4 points chosen are in general position.

Problem 27. What happens in the reconstruction problem for the L_1 or L_∞ metric?

Problem 28. If (X, ρ) is a metric space, then what are necessary and sufficient conditions on $A \subset X^2$ so that, whenever ρ' is another metric on X such that $\rho \restriction A = \rho' \restriction A$, we must have $\rho = \rho'$. What if we only want ρ and ρ' to be equivalent or uniformly equivalent?

Problem 29. Find $k(n)$ so that, if A is a metric space which can be embedded in l_∞^n, then is there a finite set $B \subset A$ of size $k(n)$ such that knowing all the distances between points of A and points of B allows one to reconstruct the distance matrix.

Problem 30. Where does L_p-embeddable fit into the scheme we have given? Does ultrametric imply L_p-embeddable which implies L_1-embeddable, when $p \neq \infty$? Are the classes of L_p-embeddable metrics comparable?

10. Hypermetric spaces and spaces of negative type

The notion of L_1-embeddable differs greatly from additive trees and ultrametrics in that it does not seem to have a definition by a simple inequality. It is suspected that there are no simple characterizations of L_1-embeddable metrics, but this has never been established.

Problem 31. Is there a first-order characterization of L_1-embeddability?

Neyman showed in 1984 [52] that there is no characterization which is a finite conjunction of inequalities. Of course, by compactness, there is a infinite conjunction of first-order formulas which characterizes L_1-embeddable.

The attempts to characterize L_1-embeddable by means of inequalities has led to some interesting inequalities which must be satisfied by any L_1-embeddable metric. These include the *hypermetric inequalities*.

Definition 5. A *hypermetric inequality* is defined for each $b : X \to \mathbb{Z}$ such that $\sum\{b(x)\colon x \in X\} = 1$ and states that $\sum\{b(x)b(y)d(x, y)\colon x, y \in X\} \leqslant 0$. A metric space which satisfies each hypermetric inequality is said to be a *hypermetric* space.

While this scheme is a little hard to understand at first, there are relatively few instances which are not satisfied automatically. In fact, the least complicated instance is accomplished by the b's which are $1, 1, 1, -1, -1$. This yields the *pentagon inequality* cited in the introduction. The easiest way to understand the hypermetric inequalities is to note that they forbid the bipartite graphs $K(n, n+1)$ when $n \geqslant 2$.

Theorem 12. *L_1-embeddable metrics are hypermetric.*

Proof. A *cut pseudometric* on a set X is a binary-valued pseudometric induced by any $A \subset X$ which is defined by letting $\rho(x, x') = 1$ iff $|\{x, x'\} \cap A| = 1$. Any L_1-embeddable metric is a linear combination of cut pseudometrics. Hypermetricity is clearly preserved by linear combinations. So it suffices to show that cut pseudometrics are hypermetric. This means that we must show that, whenever $a, b, c, d \geqslant 0$, we have

$$a + c - b - d = 1 \Rightarrow (a - b)(c - d) \leqslant 0$$

which is easy. □

Nevertheless, these inequalities do not characterize L_1-embeddable metrics. In 1977, Assouad and, independently, Avis in 1981 showed that the graph obtained by deleting two adjacent edges from K_7 is hypermetric but not L_1-embeddable. More sophisticated inequality schemes valid for L_1-embeddable metrics were devised by Deza and Laurent in 1992.

Despite their humble birth as approximations to L_1-embeddability, hypermetrics are significant to geometry. Consider the problem of identifying the metrics on $\mathbb{R}^n$ which are scalar multiples of the usual metric on each straight line (these are called projective metrics). This is Hilbert's fourth problem. In 1974, Pogorolev characterized projective metrics in $\mathbb{R}^2$. In 1986, Szabo defined a complicated example of a projective metric on $\mathbb{R}^3$ which does not satisfy Pogorolev's characterization. To see how hypermetrics are closely related to the fourth problem, we need a concept from convex geometry. A *zonoid* is a convex set which is arbitrarily close in the Hausdorff metric to convex polytopes in $\mathbb{R}^n$. Alexander showed in 1988 that whenever the dual unit ball of a finite-dimensional normed linear space M (with a projective metric) is not zonoid, Pogorolev's characterization does not work. In 1975, Kelly proved that this problem is equivalent to determining whether the dual space of M is hypermetric. To get a projective metric on $\mathbb{R}^3$ which does not satisfy Pogorolev's characterization we need only a projective metric which is not hypermetric. $L_\infty(\mathbb{R}^3)$ works!

Problem 32. Does $L_\infty(\mathbb{R}^3)$ satisfy the pentagonal inequality? Characterize the projective metrics on $\mathbb{R}^3$ which disobey the pentagonal inequality or hypermetric inequalities (or weaker properties).

It was proved in 1993 however by Deza, Grishukhin and Laurent, making use of Voronoi theory, that hypermetric spaces can be described by a finite list of inequalities. This is amazing since the hypermetric scheme is infinite and does not seem to contain any redundancies. I do not know if this follows from logical considerations alone.

Another surprising aspect of the hypermetric inequalities is that, despite their failure to characterize the L_1-embeddable metrics, they do carry some power. Indeed any hypermetric space still has some "Euclidean" structure.

Consider the example of a "distance" space consisting of the the points on the n-sphere with the metric defined by the square of the Euclidean metric. Of course, if we examine any three nearby and nearly collinear points, we see that this is not a metric space but it certainly has many metric subspaces.

Definition 6. If a metric space X can be isometrically embedded in some n-sphere with the square of the Euclidean metric, then we say that X is *spherical*.

Theorem 13 (Deza, Grishukhin and Laurent). *Every finite hypermetric space is spherical.*

Problem 33. Is any (countable, separable, arbitrary) hypermetric space isometrically embeddable in some appropriately defined κ-sphere? What is the correct infinitary notion of spherical?

Problem 34. What is an example of a spherical space which is not hypermetric?

Note that it does not suffice to take an appropriate sphere since this will not satisfy the triangle inequality.

Moving even further into weak properties, we can identify the *negative-type inequalities*. These are defined exactly like the hypermetric inequalities except that we require only $\sum\{b(x)\colon x \in X\} = 0$.

Definition 7. A *negative-type inequality* is defined for each $b : X \to \mathbb{Z}$ such that $\sum\{b(x)\colon x \in X\} = 0$ and states that $\sum\{b(x)b(y)d(x, y)\colon x, y \in X\} \leqslant 0$. A metric space which satisfies each negative-type inequality is said to be a *space of negative-type*.

Again, it is easiest to understand the negative-type inequalities as forbidding the graph $K(n, n)$ when $n \geqslant 3$. [12]

So hypermetric spaces and spaces of negative-type are defined by analogous schemes of inequalities and spherical spaces are characterized by embeddability in a specific Euclidean-style space. Nevertheless, spherical spaces interpolate hypermetric spaces and spaces of negative-type!

Theorem 14 (Deza and Grishukhin). *Every spherical space has negative-type and thus every hypermetric space has negative-type.*

Of course, metric spaces of negative-type need not be hypermetric. The graph $K(2, 3)$ demonstrates this. This graph also answers one of the two parts of the next question, but which one?

Problem 35. What is an example of a negative-type metric space which is not spherical? What is an example of a spherical space which is not hypermetric?

In the application to Hilbert's fourth problem, we used the fact that $L_\infty(\mathbb{R}^3)$ is not hypermetric.

Problem 36. Is $L_\infty(\mathbb{R}^3)$ of negative type? For which n is $L_\infty(\mathbb{R}^n)$ of negative type?

The next classical result is beautiful and surprising and demonstrates immediately why spherical metrics are of negative-type.

Theorem 15 (Schoenberg, 1938). *A metric space is of negative-type if and only if it can be embedded in some $\mathbb{R}^n$ with the metric which is the square of the Euclidean metric.*

Actually, in the language of linear algebra, this was first proved by Cayley!

Ponder Theorem 15. It says that any metric of negative type can be squared and suddenly it is embedded in Euclidean space. But this squaring is such a "nice" transformation! The reason that we have not discussed the topological level of generality since leaving additive trees becomes clear. All of these properties: L_2-embeddable, L_1-embeddable,

[12] One easily embeds $K(2, 2)$ in $\mathbb{R}^3$ and, of course, K_n is ultrametric.

hypermetric, spherical, negative type all coincide up to homeomorphism, up to uniform homeomorphism, even up to composition of the metric with a monotone function.

Let us call this composition a "scaling" and then be more exact.

Definition 8. If $f:[0,\infty)\to[0,\infty)$ is a function whose limit at zero is zero, then the scaling of a metric ρ by f is the function ρ_f defined by $\rho_f(x,y)=f(\rho(x,y))$.

Proposition 6. *Any scale which is concave up preserves the triangle inequality.*

Delistathis has noted the well-known transformation $x\to(x/(1+x))$ which is used to bound metrics provides the most common example of an application of Proposition 6.

The notion of scaling can be used to approach the problem of deciding how "geometric" these weaker metric concepts are.[13] Certainly all separable metric spaces can be embedded by a uniform homeomorphism into Hilbert space (this was proved first by Mysior, it seems). But not all separable metric spaces can be embedded by a re-scaling into Hilbert space.

Theorem 16. *There is a separable metric space which cannot be scaled to embed in a pentagonal* (*and thus, Euclidean or negative-type*) *space.*

Proof. Take the bipartite graph $K(n,n)$ for all possible choices of n and multiplied by all possible choices of positive rational numbers. □

Every finite metric space has a scale which embeds it into l_2 but whether one can get these scales in a uniform manner is unknown.

Problem 37 (Maehara, 1986). Is there a scale which embeds all metric spaces of fixed size n (even size 5) into l_2 simultaneously?

11. Lipschitz constants and eigenvalues

Another property of a transformation weaker than uniform homeomorphism but incomparable to scaling is that of an α-Lipschitz map. We say that two metrics ρ and π are α-Lipschitz where $\alpha\geqslant 1$ if every quotient $\rho(x,y)/\pi(x,y)$ and its inverse is at most α. Of course two metrics are 1-Lipschitz if and only if they coincide. This notion enables us to ask whether an arbitrary metric is α-Lipschitz to a Euclidean metric and so forth.

Note that the square root scaling is not α-Lipschitz for any constant α, so there is no reason to expect L_2-embeddable, L_1-embeddable, hypermetric, and negative-type to be α-Lipschitz for any constant α.

Proposition 7 (Bourgain, Figiel and Milman). *There is a finite metric space which is not* 2-*Lipschitz isometrically embeddable in* l_2.

[13] Note that scaling preserves ultrametricity but maybe not additive tree distances.

Theorem 17. *There is, for each $\alpha > 2$, a finite metric space which is not α-Lipschitz to a space of negative type* (*or a subset of* l_2).

Note that $K(n, n)$ is easily shown not to be $(\sqrt{2}-\varepsilon)$-Lipschitz isometrically embeddable in l_2.

Problem 38. Is there a metric space of negative type which is not α-Lipschitz isometrically embeddable to a subset of l_2?

In their pursuit of pathological examples in the geometry of Banach spaces, Bourgain, Milman and Wolfson did establish a Ramsey-theoretic theorem showing that in the disorder of arbitrary finite metric spaces can be found a certain amount of "Euclidean behavior". That is, arbitrary finite metric spaces do have fair-sized subsets which do embed into l_2.

Theorem 18 (Bourgain, Figiel and Milman). *For every $\alpha > 1$, there is $C > 0$ such that every finite metric space contains a subset which is α-Lipschitz embeddable in l_2 and has size at least $C \log |X|$.*

Indeed Bourgain, Milman and Wolfson defines their own metric inequality which says that a metric space has type 2 if there is $\varepsilon > 0$ so that, for any labeling of points by the vertices of an n-cube, the l_2-sum of the diagonals is less than ε times the l_2-sum of the edges. They show that a metric space of type 2 contains copies of l_1^n up to a Lipschitz constant.

Problem 39. Does type 2 fit naturally into the scheme of hypermetric and negative-type inequalities?

Problem 40. What Lipschitz constants, if any, exhibit the distinction between L_2-embeddable, L_1-embeddable, hypermetric, negative-type and one positive eigenvalue?

Another transformation of metrics derives from the notion of a Robinsonian metric. This is a metric ρ whose underlying set admits a linear order $\leqslant$ such that $a \leqslant b \leqslant c \leqslant d \Rightarrow \rho(a, d) \leqslant \rho(b, c)$. Thus Robinsonian metrics are metrics which are "compatible" with a linear order. Ultrametrics are Robinsonian but I know little more than this.

Problem 41. Are additive metrics Robinsonian? Are Robinsonian metrics of negative type (or hypermetric)? What if we allow $\leqslant$ to be a partial order of some kind?

Let us now turn to eigenvalues. Suppose we are given any n points in some Euclidean space and compute the distance matrix. This matrix is symmetric and thus has all real eigenvalues. It has zero entries along the diagonal and has exactly one positive and $n - 1$

negative eigenvalues. It turns out that if a metric has negative type, then it is still true that the distance matrix has exactly one positive eigenvalue.

Theorem 19. *Any metric space which is of negative type has a single positive eigenvalue.*

The existence of a single positive eigenvalue represents the weakest metric property which has so far been isolated.

Definition 9. If (X, ρ) is a metric space and, for each finite $\{a_i\colon i \in n\} \subset X$, the $n \times n$ distance matrix whose (i, j)th entry is $\rho(a_i, a_j)$ has exactly one positive eigenvalue, then we say that (X, ρ) *has one positive eigenvalue*.

To see that this definition is reasonable, one should note that if a matrix has a particular eigenvalue, then any square submatrix also has that eigenvalue. $K(3, 3)$ is not negative-type and, indeed, it has two positive eigenvalues.

Problem 42. What are the metric spaces (of smallest cardinality) which do not have one positive eigenvalue?

An example due to Winkler of a metric space with one positive eigenvalue which is not of negative type is the bipartite graph $K(5, 2)$ with a single edge added between the two points on the "side" with only two points. [14]

Problem 43. Can any metric space be scaled to have one positive eigenvalue?

The scaling method we described (taking the square root) shows that any metric of negative type can be scaled to be Euclidean, but it is unknown what happens for metrics with one positive eigenvalue.

Problem 44. Is there a metric space which has one positive eigenvalue which cannot be scaled to have negative type (equivalently, to be Euclidean)?

Problem 45. Which Tychonoff spaces have, for each continuous pseudometric, an equivalent (or generating a larger topology) continuous pseudometric with one positive eigenvalue?

Further work has been done on investigating the characteristic polynomial of distance matrices of graphs by Graham and Lovasz. This work is beyond the scope of this article, but, no doubt, investigating the characteristic polynomial of an arbitrary metric space would be rewarding.

Problem 46. Is there a useful class of metric spaces strictly weaker than those with exactly one positive eigenvalue?

[14] An elegant proof of this was given by Deza and Maehara in 1990 and Marcu in 1991.

12. Quasi-metrics

The notion of asymmetric distances occur frequently in the literature. In optimization theory, for example, the "windy postman" problem is a version of the traveling salesman problem in which the quasi-metric represents times needed to cover a distance and so, "depending on the wind", there is asymmetry.

Another significant application of asymmetric distances is in psychological measurement. The influential 1978 article by Cunningham [19] explains why this is so. "There are some situations in which the direction of the dissimilarity measurement may make a difference." He continues: "As an example, consider the case of people judging the similarity of two stimuli which differ markedly in their prominence or number of known traits". In 1977, Tversky found that people gave a consistently higher rating when asked questions like "How similar is North Korea to Red China" than when asked questions like "How similar is Red China to North Korea".

The notion of an additive tree and the notion of the four-point property both generalize to the asymmetric case naturally, but these generalizations do not seem to be equivalent. Bandelt in 1990 found equations which characterize the asymmetric generalization of additive trees.

Beside these generalizations from the symmetric case, there is no available means of classifying asymmetric distances.

The distance matrices for finite subsets of a quasi-metric spaces are not symmetric and thus these matrices may have some eigenvalues which are not real.

Problem 47. Do all quasi-metric spaces have an equivalent quasi-metric with all real eigenvalues?

Problem 48. Let X be a completely regular (topological) space. Is there, for every continuous quasi-metric on X, another continuous quasi-metric which generates a larger topology and all of whose eigenvalues are real? What if we require these quasi-metrics to generate completely regular topologies?

Problem 49. Formulate problems whose solution would make progress towards the understanding of asymmetric distance data.

13. Conclusion

The understanding of distance data is a fundamental goal of the natural and social sciences. To create this understanding, there are problems of reconstruction and approximation which are perhaps mainly problems in optimization theory and thus in linear algebra or non-linear analysis. But the problems of transformation, representation and classification are topological problems. Although the data is finite, solving the corresponding infinitary problems gives asymptotic and efficient methods for solving the finite problems.[15] Moreover, finite combinatorists find all but the most graph-theoretic of these problems far too geometric or topological.[16] Although the use of distances suggests that this is a geometric problem, the importance of transforming the data in a nonlinear manner, and the key role of approximation and reconstruction eliminates geometers from all but the most artificial and rigid of these problems. The importance of L_p in the classification may suggest that these problems lie in the territory of Banach space experts but the absence of linearity immediately disqualifies these problems from consideration by all but the most heretical of functional analysts.

This is a problem which is directly adjacent to graph theory, optimization theory, operations research, geometry, and the theory of stochastic processes. This is a problem of immediate and great importance to communications theory, to statistical mechanics, to mathematical psychology, to mathematical taxonomy and to multivariate statistical analysis whose significance will only increase when a more sophisticated theory is developed. This is a problem whose solution can be developed by topologists.

Notes added in proof

(1) Problem 1 was solved by Heikki Junnila who showed that these spaces are precisely the strongly zero-dimensional spaces.
(2) Problem 11 was solved affirmatively by J.L. Krivine and D. Dacunha-Castelle in the early 1970s.
(3) The proof of Theorem 9 in incorrect and so Theorem 9 is only a conjecture.
(4) There is now an excellent textbook by M. Deza and M. Laurent covering many of the topics of this paper. See *Geometry of Cuts and Metrics*, published by Springer-Verlag in 1997.

References

[1] R. Alexander, Zonoid theory and Hilbert's fourth problem, Geom. Dedicata 28 (1988) 199–211.
[2] M.R. Anderberg, Cluster Analysis for Applications, Academic Press, New York, 1973.
[3] P. Assouad, Sous-espaces de L^1 et inégalités hypermétriques, C. R. Acad. Sci. Paris Ser. I Math. 294 (1982) 439–442.
[4] P. Assouad, C. Delorme, Graphes plongeables dans L^1, C. R. Acad. Sci. Paris Ser. A–B 291 (1980) A369–A372.
[5] D. Avis, M. Deza, The cut cone, L^1-embeddability, complexity, and multicommodity flows, Networks 21 (1991) 595–617.
[6] D. Avis, V.P. Grishukhin, A bound on the k-gonality of facets of the hypermetric cone and related complexity problems, Comput. Geom. 2 (1993) 241–254.

[15] The importance of algorithms and complexity of computation is key to making the infinite important. If the uncountable fails, we must need enumeration and there will often be no algorithm. If the countably infinite fails, we must need to quantify over subsets and this often gives a lower bound on complexity.

[16] But it seems that a large part of the theory of distances in graphs may be extended usefully, with some work, to a theory of L_1-embeddable metrics.

[7] D.M. Avis, On the extreme rays of the metric cone, Canad. J. Math. 32 (1980) 126–144.
[8] D.M. Avis, Hypermetric spaces and the Hamming cone, Canad. J. Math. 33 (1981) 795–802.
[9] H.-J. Bandelt, Recognition of tree metrics, SIAM J. Discrete Math. 3 (1990) 1–6.
[10] J.-P. Barthélemy, A. Guénoche, Trees and Proximity Representations, John Wiley and Sons, Chichester, 1991.
[11] J.M. Bayod, The space $l^1(K)$ is not ultrametrizable, Dekker Lecture Notes in Pure Appl. Math., Vol. 137, Dekker, New York, 1992, pp. 221–255.
[12] J.M. Bayod, J. Martinez-Maurica, Subdominant ultrametrics, Proc. Amer. Math. Soc. 109 (1990) 829–834.
[13] L.M. Blumenthal, Theory and Applications of Distance Geometry, Chelsea Publishing Company, Bronx, New York, 1970.
[14] J.P. Bouchaud, P. Le Doussal, Ultrametricity transition in the graph colouring problem, Europhys. Lett. 1 (1986) 91–98.
[15] J. Bourgain, T. Figiel, V. Milman, On Hilbertian subsets of finite metric spaces, Israel J. Math. 55 (1986) 147–152.
[16] J. Bourgain, V. Milman, H. Wolfson, On type of metric spaces, Trans. Amer. Math. Soc. 294 (1986) 295–317.
[17] J.D. Carroll, Spatial, non-spatial and hybrid models for scaling, Psychometrika 41 (1976) 439–463.
[18] V.D. Chepoi, Isometric subgraphs of Hamming graphs and d-convexity, Cybernetics 24 (1988) 6–11.
[19] J.P. Cunningham, Free trees and bidirectional trees as representations of psychological distance, J. Math. Psychology 17 (1978) 165–188.
[20] M. Deza, V.P. Grishukhin, Hypermetric graphs, Quart. J. Math. Oxford Ser. 2 44 (1993) 399–433.
[21] M. Deza, M. Laurent, Facets for the cut cone, I, Math. Programming Ser. A 56 (1992) 121–160.
[22] M. Deza, M. Laurent, Facets for the cut cone, II. Clique-web inequalities, Math. Programming Ser. A 56 (1992) 161–188.
[23] M. Deza, H. Maehara, Metric transforms and Euclidean embeddings, Trans. Amer. Math. Soc. 317 (1990) 661–671.
[24] Y. Dodge (Ed.), Statistical Data Analysis, North-Holland, Amsterdam, 1987.
[25] A. Gifi, Nonlinear Multivariate Analysis, John Wiley and Sons, Chichester, 1990.
[26] A.D. Gordon, A review of hierarchical classification, J. Roy. Statist. Soc. Ser. A 150 (1982) 119–137.
[27] R.L. Graham, L. Lovasz, Distance matrix polynomials of trees, in: Theory and Applications of Graphs, Lecture Notes in Math., Vol. 642, Springer, Berlin, pp. 186–190.
[28] R.L. Graham, L. Lovasz, Distance matrix polynomials of trees, Problèmes combinatoires et théorie des graphes, in: Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976, pp. 189–190.
[29] R.L. Graham, H.O. Pollak, On the addressing problem for loop switching, Bell System Tech. J. 50 (1971) 2495–2519.
[30] R.L. Graham, H.O. Pollak, On embedding graphs in squashed cubes, in: Lecture Notes in Math., Vol. 303, 1972, pp. 99–110.
[31] R.L. Graham, P.M. Winkler, Isometric embeddings of graphs, Proc. Nat. Acad. Sci. USA 81 (1984) 7259–7260.
[32] R.L. Graham, P.M. Winkler, On isometric embeddings of graphs, Trans. Amer. Math. Soc. 288 (1985) 527–536.
[33] W. Imrich, Embedding graphs into Cartesian products, Ann. New York Acad. Sci. 576 (1986) 266–274.
[34] N. Jardine, R. Sibson, Mathematical Taxonomy, John Wiley and Sons, London, 1971.
[35] J.B. Kelly, Hypermetric spaces, in: Lecture Notes in Math., Vol. 490, Springer, Berlin, 1975, pp. 17–31.

[36] M. Krivanek, A note on the computational complexity of hierarchical overlapping clustering, Aplikace Mat. 30 (1985) 453–460.
[37] M. Krivanek, The complexity of ultrametric partitions on graphs, Inform. Process. Lett. 27 (1988) 265–270.
[38] P. Krivka, N. Trinajstic, On the distance polynomial of a graph, Aplikace Mat. 28 (1983) 357–363.
[39] A.Y. Lemin, Isometric imbedding of isosceles (non-Archimedean) spaces into Euclidean ones, Dokl. Akad. Nauk SSSR 285 (1985) 558–562.
[40] W.-H. Li, Simple method for constructing phylogenetic trees from distance matrices, Proc. Nat. Acad. Sci. USA 78 (1981) 1085–1089.
[41] H. Maehara, Metric transforms of finite spaces and connected graphs, Discrete Math. 61 (1986) 235–246.
[42] S.M. Malitz, J.I. Malitz, A bounded compactness theorem for L^1-embeddability of metric spaces in the plane, Discrete Comput. Geom. 8 (1992) 373–385.
[43] D. Marcu, Note on the Euclidean embeddings, Quart. J. Math. Oxford Ser. 2 42 (1991) 87–98.
[44] M. Marcus, T.R. Smith, A note on the determinants and eigenvalues of distance matrices, Linear and Multilinear Algebra 25 (1989) 219–230.
[45] R. Mathar, The best Euclidian fit to a given distance matrix in prescribed dimensions, Linear Algebra Appl. 67 (1985) 1–6.
[46] R. Mathar, Multidimensional scaling with constraints on the configuration, J. Multivariate Anal. 33 (1990) 151–156.
[47] J.C. Mayer, L.K. Mohler, L.G. Oversteegen, E.D. Tymchatyn, Characterization of separable metric $\mathbb{R}$-trees, Proc. Amer. Math. Soc. 115 (1992) 257–264.
[48] J.C. Mayer, J. Nikiel, L.G. Oversteegen, Universal spaces for $\mathbb{R}$-trees, Trans. Amer. Math. Soc. 334 (1992) 411–432.
[49] J.K. Misiewicz, Positive definite norm dependent functions on l^∞, Statist. Probab. Lett. 8 (1989) 255–260.
[50] H. Movahedi-Lankarani, On the inverse of Mane's projection, Proc. Amer. Math. Soc. 116 (1992) 555–560.
[51] W.C. Navidi, G.A. Churchill, A. von Haeseler, Phylogenetic inference: linear invariants and maximum likelihood, Biometrics 49 (1993) 543–555.
[52] A. Neyman, Representation of L_p-norms and isometric embedding in L_p-spaces, Israel J. Math. 48 (1984) 129–138.
[53] D. Penny, Towards a basis for classification: the incompleteness of distance measures, incompatibility analysis and phenetic classification, J. Theoret. Biology 96 (1982) 129–142.
[54] D. Penny, L.R. Foulds, L.M. Hendy, Testing the theory of evolution by comparing phylogenetic trees constructed from five different protein sequences, Nature 297 (1982) 197–200.
[55] R. Rammal, G. Toulouse, M.A. Virasoro, Ultrametricity for physicists, Rev. Modern Phys. 58 (1986) 765–788.
[56] K. Rudnik, Concerning the rigidity problem for subsets of E^2, Bull. Polish Acad. Sci. Math. 37 (1989) 251–254.
[57] S. Sattath, A. Tversky, Additive similarity trees, Psychometrika 42 (1977) 319–345.
[58] J. Schlitter, Calculation of coordinates from incomplete and incorrect distance data, Z. Angew. Math. Phys. 38 (1987) 1–9.
[59] R.B. Shepard, Multidimensional scaling, tree-fitting, and clustering, Science 210 (1980) 390–398.
[60] M.J. Sippl, H.A. Scheraga, Cayley–Menger coordinates, Proc. Nat. Acad. Sci. USA 83 (1986) 2283–2287.
[61] P. Sneath, R. Sokal, Numerical Taxonomy, W.H. Freeman and Co., San Francisco, CA, 1973.
[62] A.F. Timan, I.A. Vestfrid, Any separable ultrametric space is isometrically embeddable in l_2, Funk. Anal. i Prilozhen. 17 (1983) 85–86.

[63] B. Van Cutsem, Minimal superior ultrametrics under order constraint, Discrete Math. 84 (1990) 201–204.
[64] I.A. Vestfrid, A.F. Timan, A universality property of Hilbert spaces, Dokl. Akad. Nauk SSSR 246 (1979) 528–530.
[65] E.O. Wiley, Phylogenetics, John Wiley and Sons, New York, 1981.
[66] P.M. Winkler, Proof of the squashed cube conjecture, Combinatorica 3 (1983) 135–139.
[67] H.S. Witsenhausen, Minimum dimension embedding of finite metric spaces, J. Combin. Theory Ser. A 42 (1986) 184–199.
[68] L. Yang, J.-Z. Zhang, Metric spaces which cannot be isometrically embedded in Hilbert space, Bull. Austral. Math. Soc. 30 (1984) 161–167.